STUDIA PHONOLOGICA III (1964)

Experimental Observation on the Noisy Quality of Hoarseness

Naoaki YANAGIHARA

Introduction

There has been a considerable interest in recent years in the description of the physiological and acoustical events which give rise to listeners' perceptions of hoarse or harsh voice quality. Several investigators have made acoustical analysis of hoarseness by means of soundspectrographic method, and have called attention to the additional noise components which were notable features in any cases of hoarse voice. The auther made detailed soundspectrographic studies on thirty cases of hoarseness previously with special reference to the pathologic change of the frequency pattern, and the pattern of additional noise component. As the results, following basic views were described.

1) Additional noise component appears in the main formant range of every vowel (in Japanese u o a e i) in the cases of slight degree of hoarseness.

2) With the progress of the deterioration of voice, pattern of the higher frequency components becomes irregular and the replacement of the higher frequency pattern by the additional noise component takes place from the higher frequency toward the lower.

3) Distribution of additional noise component differs considerably from vowel to vowel. Generally, additional noise component is much more evident in the vowels "e" and "i" than in vowels "u'' and "o".

Based on these findings observed, soundspectrogram of hoarseness is classified into four cardinal types.

Type 1: Additional noise component extists in the main formant range of every vowel and frequency component maintains almost normal property.

Type 2: Poor regularity of frequency pattern and increment of noise energy in the second formant range of "e" and "i" are notable in addition to the finding of type one. In some cases weak noise pattern is recognized in the high frequency range about 3 Kc or 5 Kc.

Type 3: The frequency patterns of the second formant of "e" and "i" are completely replaced by noise, and noise in higher frequency range becomes more intense. Range of this higher frequency noise expands obviously in the vowels "i", "e" and "a", intensifying its energy. Even in this type, noise patterns in vowels "u" and "o" are confined within low frequency around the first formant.

Naoaki YANAGIHARA (柳原尚明) : Department of Otolaryngology, Faculty of Medicine Kyoto University

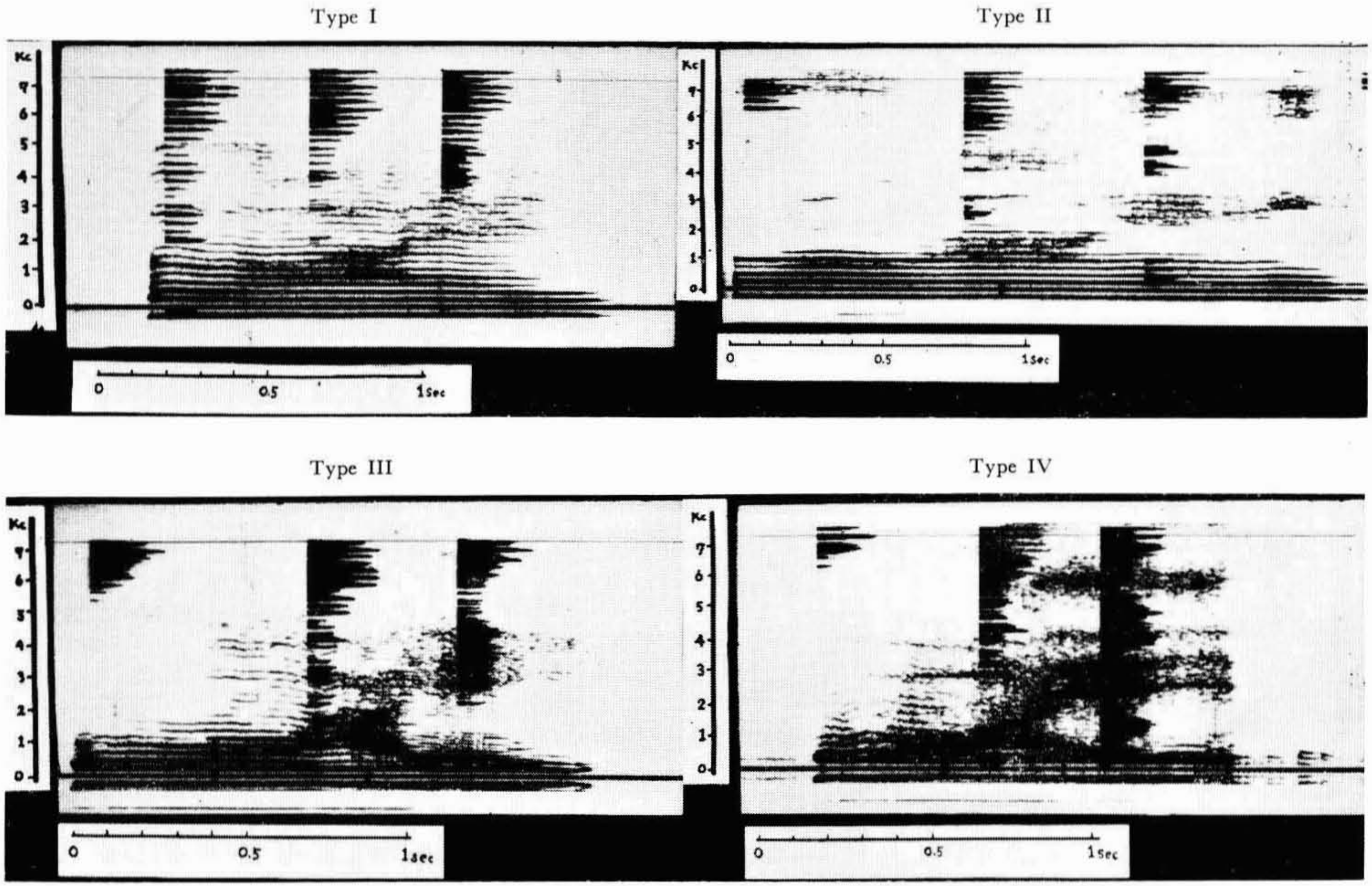


Fig. 1. Representative soundspectrogram pattern of every type of hoarseness.

Type 4: Strong energy of noise in the second formant range of "i", "e" and "a", is characteristic. Conspicuous noise pattern is often notable even in vowels "u" and "o". The frequency patterns of "i", "e" and "a" are almost completely replaced by noise components. Higher frequency noise and its expansion is much more remarkable in this type, especially vowels "i", "e" and "a".

Representative soundspectrogram of each type is illustrated in Fig. 1, and schematicaly drawn in Fig. 2, which will be helpful for better understanding of

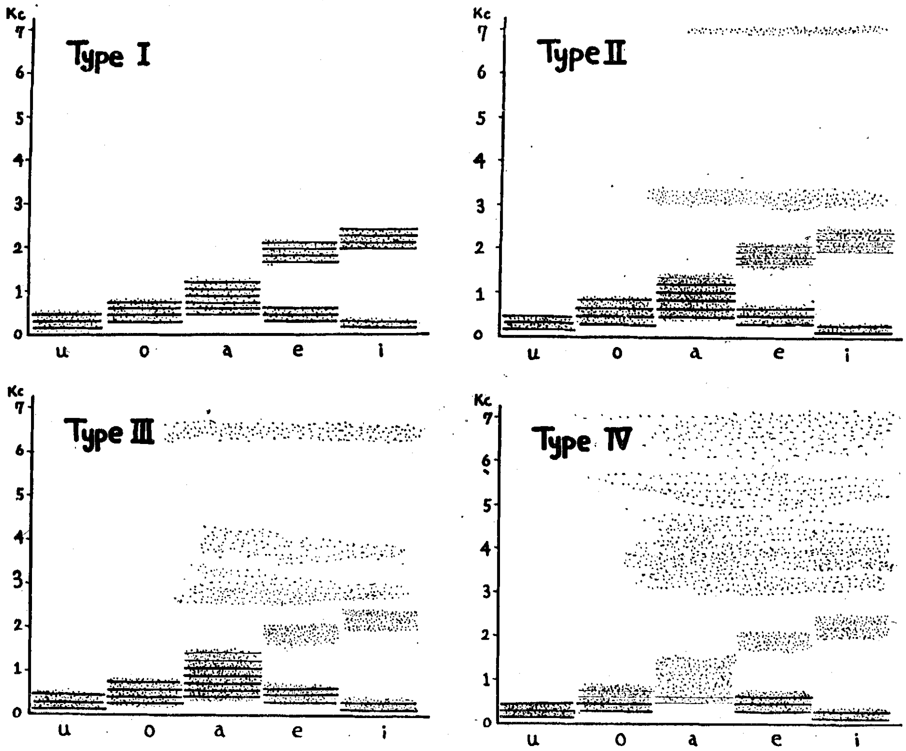


Fig. 2. Schematic drawing of soundspectrogram of every type of hoarseness.

this classification. It is reasonable to presume that listener's perception of hoarse quality in voice mainly depends on the intensity and distribution of additional noise components and the reduction of harmonic components of the vowels.

To test the validity of this conception, an attempt was made to synthesize hoarse vowels using vowel-band noise materials. The study to be reported here will deal with the correlation between the degree of hoarseness and the acoustical structure of hoarseness.

Experimental Procedure

Synthesis of hoarse voice

An outline of the synthetic method of hoarse voice is presented in the simplified block diagram of Fig. 3. As the materials, both white noise produced by

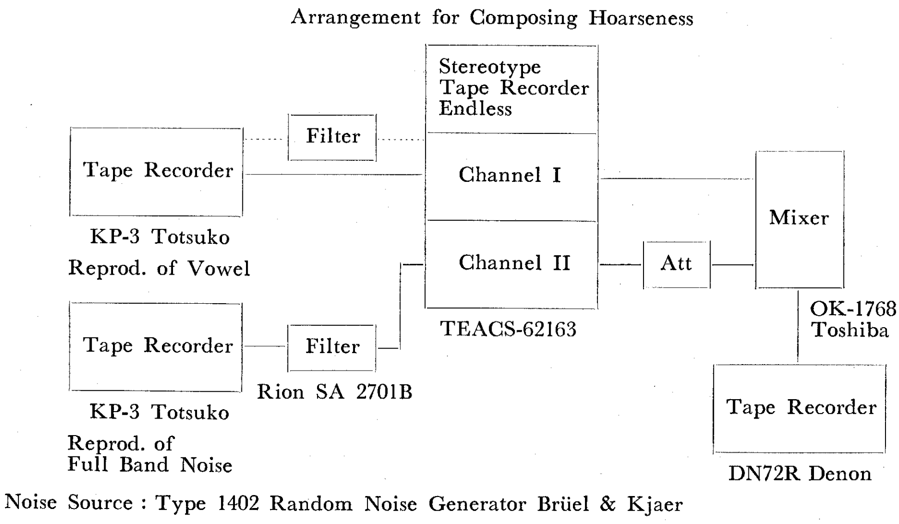


Fig. 3. Symplified block diagram of synthetic method of hoarseness.

the random noise generator type 1402 (Brüel & Kjær) and sustained vowels "a" and "e" spoken by a professional male anouncer were used. These vowel materials were picked up by an undirectional condenser microphone (CN-1 Totsuko) which was placed 20 cm from the mouth and recorded on full track tape recordor (KP-3 Totsuko) at the level of 0db. The voice was quite clear in quality, and its pitch was kept at 140 cps. A vowel material was reproduced and recorded directly or after low pass filtering on the one channel of the stereotype tape recorder (TEAC CS-62163), which had an endless reproduction apparatus. On the other channel of this stereotype tape recorder, the band filtered random noise was recorded parallel with the vowel material. Both vowel-band noise material reproduced endlessly. The reproducing level of the vowel was adjusted constant at 0db of VU meter. While the reproducting level of the band noise materials was varied in accordance with the position of the filtering band and the vowel material. Through a mixer (OK-1768 Toshiba), the outputs of both channels of the stereotype tape recorder was mixed and put into the line input of a full track tape recorder (Denon 72R) and recorded at the constant level of 0db of VU meter.

Thus, the vowel-band noise test sounds were made and cut into short segments for arrangement in random order. The duration of test sound was four seconds, and an interval of four seconds between two successive test sounds was adopted.

Measurement of the degree of hoarseness

Six trained laryngologists who were the staff members of Dept. of otolaryngology faculty of medicine Kyoto Univ., served as listeners for this experiment. None had a history of hearing loss. Each listener was instructed to listen to a series of test sound and attempt to give one of the following six ratings to every test sound when test sound was heard as natural quality of hoarseness. When test sound was heard unnatural, instruction was given to check as unnatural.

Six ratings showing the degree of hoarseness were: (0) normal voice, (1) very slightly hoarse, (2) slightly hoarse, (3) moderately hoarse, (4) highly hoarse, and (5) extremely hoarse.

Degree of hoarseness was expressed by a mathematical average value of the total rating numbers given by six listeners.

Results

Soundspectrograms of all the test materials were made for the purpose of visualizing the situation of synthesis. This visualization of the acoustic structure of every test sound served to understand the correlation between acoustic structure of sound and corresponding listeners' perceptions of hoarseness. Above all, this procedure is of special value when the experimental result is compared with the soundspectrogram of actual hoarseness.

Table I and II shows the results of this experiment. The synthetic sounds listed in these tables were heard as natural hoarseness by all the listeners. In order to reach effective understanding, sound spectrograms of important value are also illustrated.

Items 1, 2, 3 and 4, and 1, 2 and 3 in tables I and II respectively show the significance of additional noise components in major formant ranges. Figs. 2 and 3 clearly show this significance. When the band noises were added in the second and third formant ranges, remarkable increase of mean of rating is recognized. Item 4, 5, 6, and 7, 8, 9, in Table II, also denote the significance of noises in the second formant range.

Table I. Structure of synthetic sound and corresponding mean of rating. Vowel "a".

	Vowel + Band Noise	Mean of Rat.
1	Vowel "a"	0.3
2	Vowel "a"+609∼1700	1.8
3	Vowel "a"+600∼1700+2400∼3400	3.8
4	Vowel "a"+600∼1700+3400∼4800	3.0
5	1200 cps LP "a"+600∼1700	4.1
6	1200 LP "a"+600∼1700+2400∼3400	4.5
7	1200 LP "a"+600∼1700+3400∼4800	4.5
8	1200 LP "a,'+600∼1700+3400 LP	4.8
9	1200 LP "a"+600∼1700+4800 LP	4.3

Relative Intensity

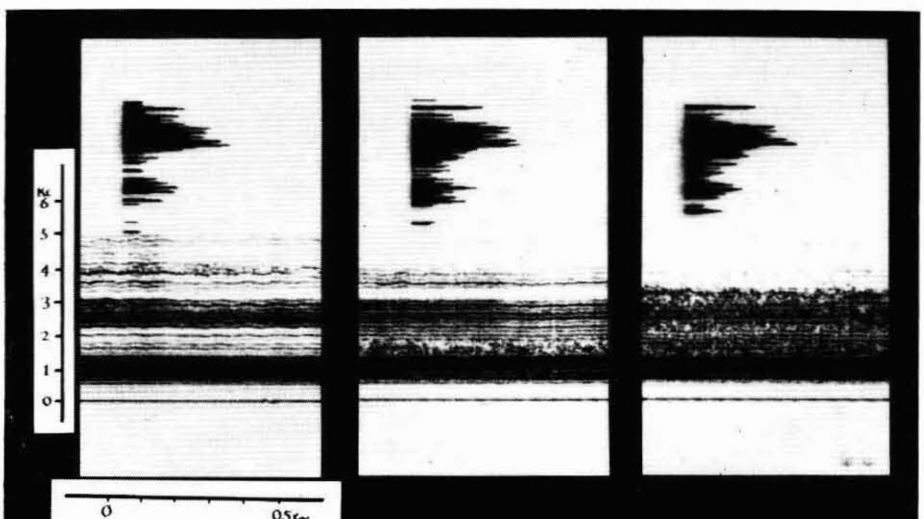
Vowel and Low Pass Vowel	0db
600∼1700 cps Band Noise	−10db
2400∼3400 Band Noise	−25db
3400∼4800	
Low Pass Noise	−25db

Table II. Structure of synthetic sound and corresponding mean of rating. Vowel "e".

	Vowel + Band Noise	Mean of Rat.
1	Vowel "e"	0.2
2	Vowel "e" + 300～6000 cps	1.1
3	Vowel "e" + 300～600 + 1700～2400	2.2
4	2400 cps LP "e" + 300～600	1.1
5	2400 cps LP "e" + 300～600 + 1700～2400	2.8
6	2400 cps Lp "e" + 300～600 + 1700～3400	2.5
7	1700 cps LP "e" + 300～600	2.5
8	1700 cps LP "e" + 300～600 + 1700～2400	3.2
9	1700 cps LP "e" + 300～600 + 1700～2400	3.5
10	1700 LP "e" + 300～600 + 1200～2400 + 3400 LP	3.8
11	1700 LP "e" + 300～600 + 1200～2400 + 4800 LP	4.3
12	1700 LP "e" + 300～600 + 1700～3400 + 4800 LP	4.4
13	1700 LP "e" + 300～600 + 1700～3400 + 6800 LP	4.4

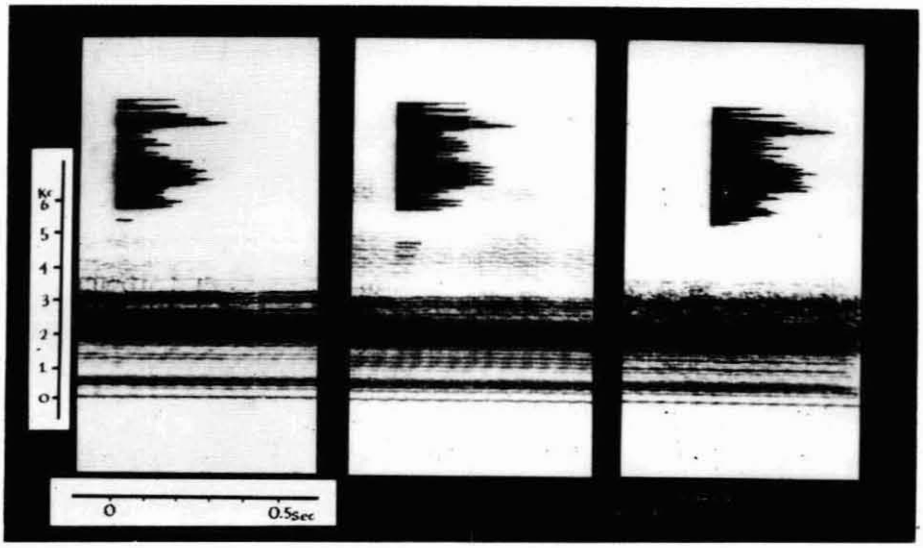
Relative Intensity

Vowel and Low Pass Vowel	0db
300～600 cps Band Noise	−25db
1700～2400	
700～2400 Band Noise	−25db
1200～2400	
Low Pass Noise	−30db



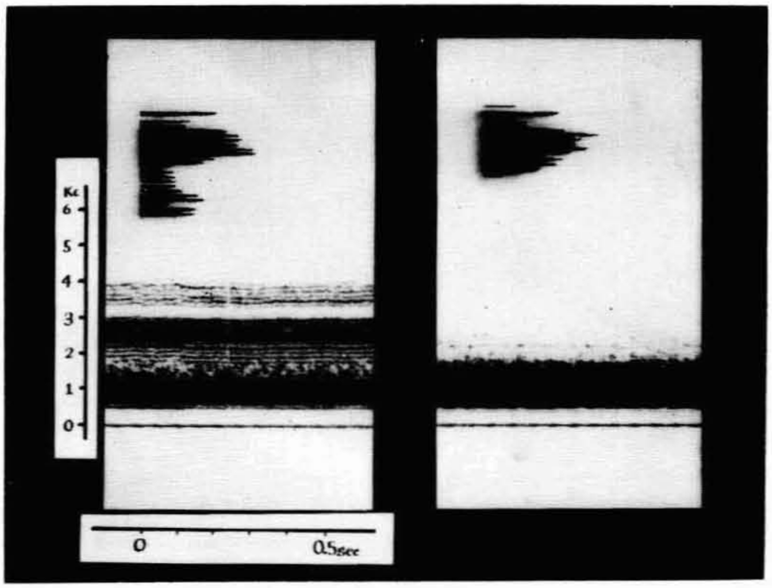
Vowel	"a"	"a"	"a"
Band Noise		600～1700 cps	600～1700 cps 2400～3400 cps
Mean of Rating	0.3	1.8	3.8

Fig. 4. Soundspectrogram of vowel "a" plus band noise test sound showing the status of synthesis. Bandwidth of filter and mean of rating indicated under corresponding sound.



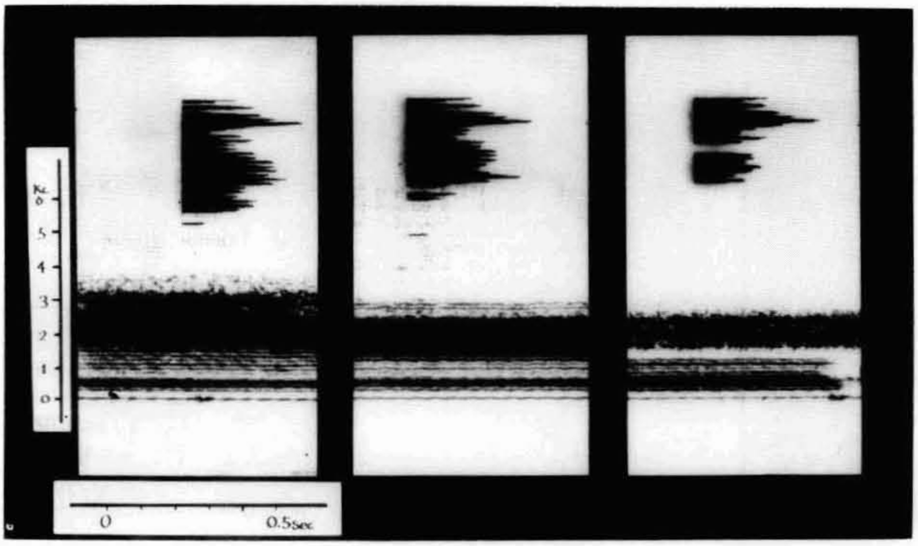
Vowel	"e"	"e"	"e"
Band noise		300～600 cps	300～600 cps 1700～2400 cps
Mean of Rating	0.2	1.1	2.2

Fig. 5. Soundspectrogram of vowels "e" plus band noise test sound showing the status of synthesis. Bandwidth of filter and mean of rating indicated under corresponding sound.



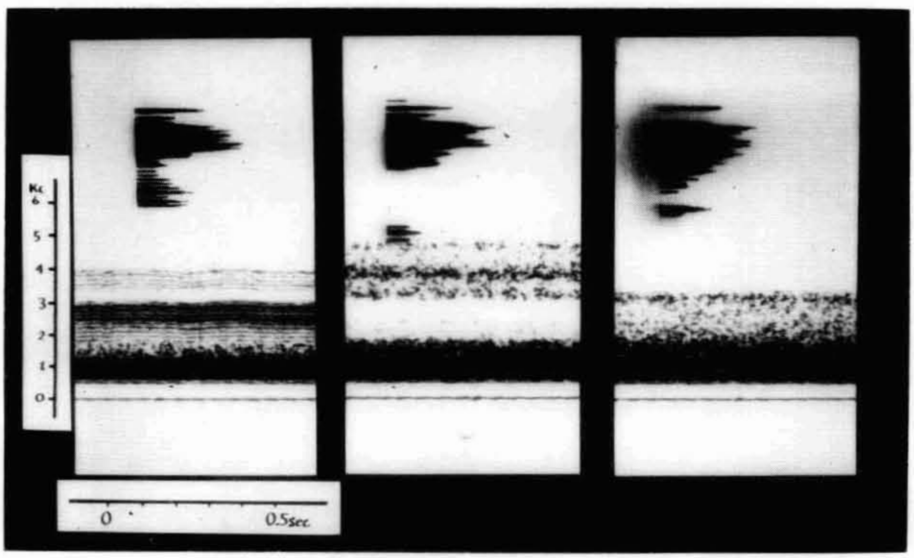
Vowel	"a"	1200 cps Low Pass "a"
Band Noise	600～1700 cps	600～1700 cps
Mean of Rating	1.8	4.1

Fig. 6. Sound spectrogram showing the effect of lowpass filtering of vowel "a".



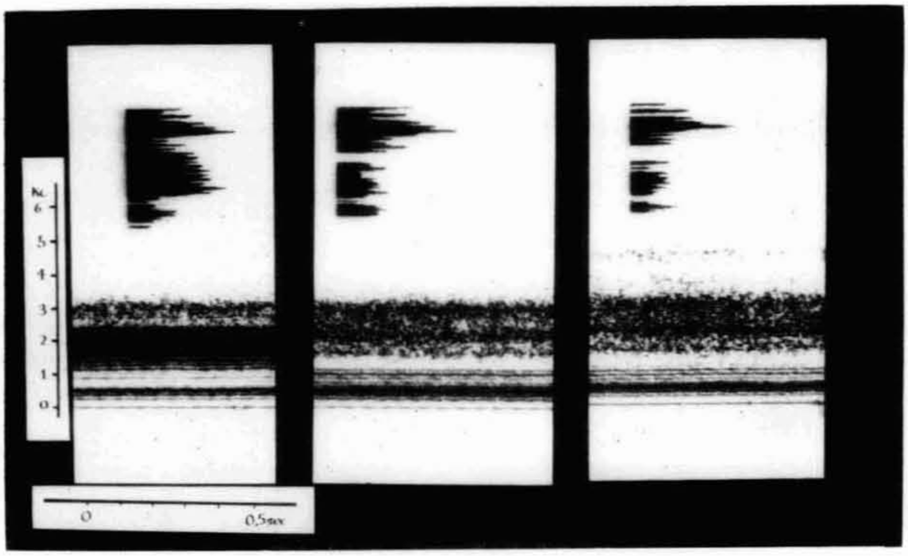
Vowel	"e"	2400 cps Low Pass "e"	1700 cps Low Pass "e"
Band Noise	300～600 cps 1700～2400 cps	300～600 cps 1700～2400 cps	300～600 cps 1700～2400 cps
Mean of Rating	2.2	2.8	3.2

Fig. 7. Soundspectgram showing the effect of lowpass filtering of vowel "e". Note step by step increase of mean of rating.



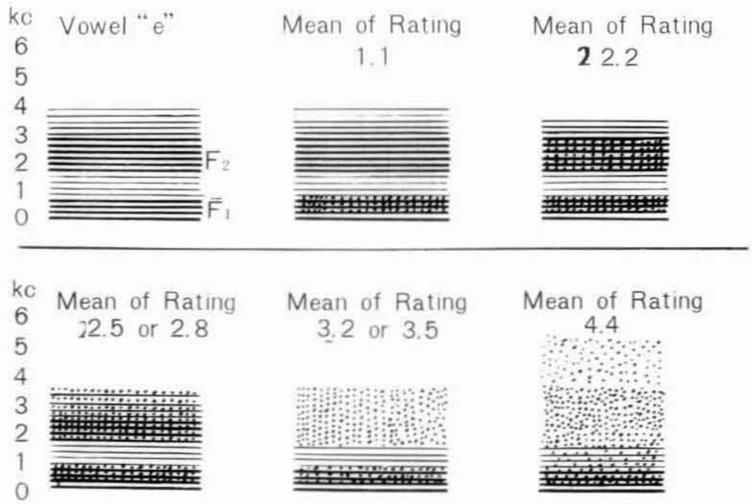
Vowel	"a"	1200 cps Low Pass "a"	1200 cps Low Pass "a"
Band Noise	600～1700 cps	600～1700 cps 3400～4800 cps	600～1700 cps 3400 cps Low Pass
Mean of Rating	1.8	4.5	4.8

Fig. 8. Soundspectrogram showing the effect of expansion of additional noise. Vowel "a".



Vowel	2400 cps Low Pass "e"	1700 cps Low Pass "e"	1700 cps Low Pass "e"
Band Noise	300～600 cps 1700～3400 cps	300～600 cps 1700～3400 cps	300～600 cps 1700～3400 cps 4800cps Low Pass
Mean of Rating	2.5	3.5	4.4

Fig. 9. Soundspectrogram showing the effect of expansion of additional noise. Vowel "e".



Schematic Presentation of Soundspectrogram of Composed Hoarseness and its Mean of Rating

Fig. 10. Schematic presentation of soundspectrogram of synthetized hoarseness and its mean of rating.

Increments of mean of ratings were resulted by the elimination of high frequency component of vowel. Items 5, 6, and 7 in Table I show the effect of low pass filtering of vowel "a". It can be seen in Fig. 4, how this elimination is effective to give hoarse sensation to the test sounds. Items 5, 6, 7, and 7, 8, 9 in Table II show the result of high frequency elimination in the case of vowel "e". In figure 2, the effect of this elimination is clearly seen. The more the energy of high frequency component was eliminated, the more increasing of mean of rating was noted. Because bandwidth of the filtered noise and, the levels of both vowels and noises kept constant, the increment of mean of rating could be attributable to the effect of high frequency elimination of vowel materials.

In a most advanced case of hoarseness, expansion of additional noise component is an outstanding feature. To imitate this situation, 3400 cps, 4800 cps and 6800 cps low pass noise were added to the low pass vowel-band noise test sounds. The mean of rating markedly increased in the vowel "e", while in the vowel a more or less slightly increasing of mean of rating is observed. Step by step increasing of mean rating and its mate of test sound is schematically presented in Fig 8.

DISCUSSION

There are two different opinions concerning what the harsh voice or hoarse voice is and how the acoustic structure of the deteriorated voice is related to malfunction of the vocal cord. The one opinion is that harsh voice quality is related to aperiodicity of fundamental frequency of voice. The other is that harsh voice quality of hoarse voice quality is related to additional noise component observed in a sound spectrogram of voice. According to Nessel's view, reduction of higher harmonic component and additional noise component of high frequency modulated by formant are notable in hoarseness. In most cases of hoarseness, additional noise component above 5 kc is marked characteristically, which is independent from the fundamental frequency of voice and the species of vowel. Generally noise element is more remarkable in the vowels "e" and "i". There is a remarkable agreement between this view and the auther's conclusion described previously. These acoustical findings obtained by analysis raised a question as to whether the acoustical finding related to the degree of hoarseness and to malfunction the vocal cords.

By means of synthetic method, process of the deterioration of hoarse voice, and method of evaluating the degree of hoarseness were investigated. The most essential factors which give hoarse sensation to voice are the additional noise components. It can be said that the status of hoarse voice is determined by these factors recognizable in the soundspectrogram of voice.

Yet little is known about interrelationship between acoustical pattern of the noisy quality of hoarse voice and aperiodic vibratory function of the vocal cords. Having instrumental limitation, the soundspectrographic method can not provide information on the aperiodicity of fundamental frequency, cycle to cycle fundamental variation. By means of electrical laryngeal analog Wendahl suggested that cycle to cycle fundamental variation had much to do with harsh voice quality. At the median frequency of 100 cps, one cycle of variation could give harsh quality to voice. Lieberman observed the perturbation of fundamental frequency of speech in the cases of normal and pathologic larynx. According to his accurate measurement, perturbation over 0.5 msec is attributed to the abnormal variation of vibratory pattern of the vocal cords. He suggested that the acoustical method would be of practical value as a screening method of laryngeal disease especially laryngeal malignancy.

In order to know full significance of acoustic features of pathologic voice, fundamental frequency variation, reduction of harmonic component, and additional noise component, are to be considered altogether.

Summary

The purpose of the present investigation was to explore the relationship between the acoustical findings obtained by the sound spectrographic analysis of hoarseness and the degree of hoarseness. Synthesis of hoarseness was attempted in order to know the factors which are concerned with quality of voice. The method of synthesis is described and the degree of hoarseness of synthetic sound was measured by rating method.

On the basis of the experiment, the following conclusions were drawn.

1. To give hoarse sensation of voice, noise components must be added to the main formant ranges of vowels, among which the second and third formant ranges paly major role.

2.· The factors related to the degree of hoarseness are high frequency additional noise, reduction of high frequency harmonic component, and expansion of the range of additional noise components. The status and degree of hoarseness is determined by these factors recorgnized in the soundspectrogram.

References

1. Wendahl, R. W. : Laryngeal Analog Synthesis of Harsh Voice Quality. Folia Phoniat. 15: 241 : 1963
2. Schönhärl, E. : Beitrag zur Qaulitativen Stimmanalyse. Z. Laryng. Rhinol. 42 : 13 : 1963
4. Lieberman, P. : Some Acoustic Measures of the Fundamental Periodicity of Normal and Pathologic Larynges. J.A.S.A 35 : 334 : 1963
5. Nessel, E. : Uber das Tonfrequenzspektrum der Pathologische Veränderten Stimme. Acta oto-laryng. Suppl. 157 ; 1960
6. Nishikawa, T. : Study on the Relationship between the Voice and the Status of the Vocal Cord. J. of the Otorhino-Laryng. Society of Japan, 65 : 544 : 1962
7. Yanagihara, N. : Acoustical Studies on Hoarseness. Zibi-rinsyo, 55 : 357 : 1962
8. Takakura, M. : The Quantitative Measurement of Degree of Purity or of Hoarseness of Human Voice. J. of the Otorhino-Laryng. Society of Japan, 63 : 1259 : 1960